1 gatategeca ecatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga md amkr qlc cvl llcq 61 gcagtetteg tttegeccag cagegetggg ccaactgtge cegacagaga caatgatgga vspssag ptv pdr dndg 121 atccctgata gtctagaggt tgagggatac acggtagatg tcaagaacaa aaggactttt ipd sle vegy tvd vkn krtf 181 ctctcgcctt ggatctcaaa tatccatgag aagaagggc ttaccaaqta caaqtcctcc lsp wis nihe kkg ltk ykss 241 cccqaqaaqt qgtctaccgc ttccqatcca tataqcqatt tcqaqaaqgt cacaqqccqq pek wstasdp ysd fek 301 ategataaaa atgtgtetee agaggetaga caceccetgg tagcageeta eeegattgta idk nvs pear hpl vaa 361 cacqtggaca tggagaacat cattctaagc aaaaacgagg accagtccac acaaaacact hvd men iils kne das 421 gacteegaga ceegcaccat atetaaaaac accagtaett caaggaccca cacetetgaa dsetrtiskntstsrt 481 qtqcacqqca atqcggaagt ccatgcatcg tttttcgata ttqqtqqctc cqtqtcaqcc vhq nae vhas ffdiqqs vsa 541 ggctttagca atagcaactc ctcgacggtt gccattgacc actcactgtc attagcaggt qfs nsn sstvaid hsl slag 601 gagaggaett gggetgaaac tatgggtetg aatacegeeg atacggeeeg geteaaegea ert wae tmgl ntadta 661 aatatteggt aegteaacae agggaetget eetatatata aegtgetgee taegacaagt nir yvn tgta piy nvl cttgtcctgg gcaaaaatca gaccctcgca accattaagg caaaggaaaa tcagctgagc lvlqknqtlatikake 781 cagatecteg ecectaacaa etattateca tecaaaaatt tageeeccat ageeetgaac qilapn nyypskn lapialn 841 geocaggaeg actititecte tacceccata actatgaatt acaatcagti ectqqaqetg aqd dfs stpitmn ynq flel 901 gaaaagacga agcagctgag actagacacc gatcaggtgt atggaaacat agcgacatat ekt k q l r l d t d q v y q n

961 aactttgaga acggcegegt gegegtegae actgggteaa actggtetga agttetgeeg
n f e n g r v r v d t g s n w s e v l p

1021 caaattcaag agacaacege cagaattate tttaatggga aggacttgaa cettgtegaa
g i q e t t a r i i f n g k d l n l v e

- 1081 egtagaattg cegeegtgaa eeceagtgat eeaetegaga egaetaaace ggatatgaca r r i a a a v n p s d p l e t t k p d m t
- 1141 ctgaaagagg ctctgaagat tgccttcgga ttcaacgaac ctaatggcaa tttgcagtat 1 k e a l k i a f g f n e p n q n l q y
- 1201 caggggaaag acatcacaga gtttgatttc aatttcgatc agcagacttc ccaaaatatc q g k d i t e f d f n f d q q t s q n i
- 1261 aaaaatcagt tggcagagct gaatgccacc aatatctaca cggttctcga taaaatcaaa k n q l a e l n a t n i y t v l d k i k
- 1321 cttaacgcca agatgaacat attgattcga gacaaacgct tccactacga ccgcaacaat
 l n a k m n i l i r d k r f h y d r n n
- 1381 atagccgtag gcgctgatga gtctgtcgtc aaggaggctc atagggaagt tatcaacagc i a v g a d e s v v k e a h r e v i n s
- 1441 agtactgaag ggctgttact taatatcgac aaggacattc ggaagatcct gtccgggtat
 s t e q l l l n i d k d i r k i l s q y
- 1501 atcgtggaga tcgaggatac cgagggcctg aaggaagtca ttaacgaccg ctatgatatg
 i v e i e d t e g l k e v i n d r v d m
- 1561 ctgaacattt ccagcttacg acaggacggt aagacattta ttgactttaa aaagtataac l n i s s l r q d g k t f i d f k k y n
- 1681 aaggagaaca caatcatcaa tccaagcgag aacggcgata ccagcacaaa tggaatcaaa k e n t i i n p s e n g d t s t n g i k
- 1741 aagatcctta tatttagtaa aaaaggctac gagatcggtt gaggatcc k i l i f s k k g γ e i g -

1 gatategeca ccatggatge aatgaagaga gggetetget gtgtgetget getgtgtgqa geaqtetteg md amkr glc cvl llcq avf 71 .tttcgcccag cagcgctggg ccaactgtgc ccgacagaga caatgatgga atccctgata gtctagaggt vspssagpt v pdr dndg i pd s 1 e 141 tgagggatac acggtagatg tcaagaacaa aaggactttt ctctcgcctt ggatctcaaa tatccatgag vegy tvď vkn krtflsp wis nihe 211 aagaaggggc ttaccaagta caagtcctcc cccgagaagt ggtctaccgc ttccgatcca tatagcgatt kkqltk yksspek wstasdp vsd 281 tcgagaaggt cacaggccgg atcgataaaa atgtgtctcc agaggctaga caccccctgg tagcagccta fek vtgridk nvs pear h p l v a a 351 cccgattgta cacgtggaca tggagaacat cattctaagc aaaaacgagg accagtccac acaaaacact ypivh vd men iils kne dqs tqn t 421 gacteegaga eeegcaccat atetaaaaac accagtaett caaggaccca cacetetgaa gtgeaeggea dsetrtiskn tst srt htse v h q 491 atgcggaagt ccatgcatcg gatattggtg gctccgtgtc agccggcttt agcaatagca actcctcgac nae vhas dig gsv sagf sns nss 561 ggttgccatt gaccactcac tgtcattagc aggtgagagg acttgggctg aaactatggg tctgaatacc t vaidh s 1 s 1 ager twaetm q 1 n t 631 gccgatacgg cccggctcaa cgcaaatatt cggtacgtca acacagggac tgctcctata tataacgtgc adtarl naniry v ntgtapiyn v 701 tgcctacgac aagtettgte etgggcaaaa atcagaceet egcaaceatt aaggcaaagg aaaatcaget 1 pt tslv l q k n q t l a t i k a k e n q 771 gagecagate etegeceeta acaactatta tecatecaaa aatttageee eeatageeet gaacgeeeag lsqi lap nny ypsk nlapia lnaq 841 gacgactttt cctctacccc cataactatg aattacaatc agttcctgga gctggaaaag acgaagcagc ddf s s t p i t m n y n q f l e l e k t k q 911 tgagactaga caccgatcag gtgtatggaa acatagcgac atataacttt gagaacggcc gcgtgcgcgt 1 r 1 d t d q v y q n i a t v n f e n q r v r 981 cgacactggg tcaaactggt ctgaagttct gccgcaaatt caagagacaa ccgccagaat tatctttaat vdtg snw sevlpqi qet tar i i f n 1051 gggaaggact tgaaccttgt cgaacgtaga attgccgccg tgaaccccag tgatccactc gagacgacta gkd lnl verriaa v n p sd plett 1121 aaccggatat gacactgaaa gaggetetga agattgeett eggatteaac gaacetaatg geaatttgea kpd mtlkeal kia fgfnepngn1 1191 gtatcagggg aaagacatca cagagtttga tttcaatttc gatcagcaga cttcccaaaa tatcaaaaat qyqg kdi tefdfnfdqq tsq nikn 1261 cagttggcag agctgaatgc caccaatatc tacacggttc tcgataaaat caaacttaac gccaagatga qlaelnatniytvldkikinak m 1331 acatattgat tegagacaaa egetteeact aegacegeaa caatatagee gtaggegetg atgagtetgt n'ilirdkrfhydr nnia v gades 1401 cgtcaaggag getcataggg aagttateaa cageagtaet gaagggetgt taettaatat egacaaggae v v k e ahr evinssteglllnidkd 1471 attcggaaga tcctgtccgg gtatatcgtg gagatcgagg ataccgaggg cctgaaggaa gtcattaacg irkils gyive ied teglke vin

Figure 2B

- 1541 accgctatga tatgctgaac atttccagct tacgacagga cggtaagaca tttattgact ttaaaaagta d r y d m l n i s s l r q d g k t f i d f k k

}

1 gatategeca ccatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga md amkr glc cvl llcg 61 gcagtctteg tttcgcccag cgaagtgaag caagaaaatc gacttctgaa cgagagcgaa avf vsp sevk qen r 1 1 121 agttcatcac agggtettet eggatactac ttcagtgact tgaattteca agcaccaatg sss qgllgyy fsdlnf qapm 181 qtqqtqacta qtaqcaccac cqqcqatttq aqcattecca getetqaqtt qqaqaacatt vvt sst tgdl sip sse 241 cccagcgaaa atcagtactt ccagtctgct atctggtccg gattcattaa ggttaaaaag pse nqy fqsa iws gfi kvkk 301 tecgaegaat atacatttge taceteggeg gataaceatg tgacaatgtg ggtggaegae sde y t f atsadnh v t m w v d d 361 caggaagtga tcaacaaggc ttcaaactct aataaaatcc ggctcgagaa ggggaggctc qevink asns nki rle kgrl 421 taccagatca aaattcagta ccagcgggaa aaccctacag aaaaaggact cgatttcaag yqi kiq yqre nptekgldfk 481 ctgtactqqa caqatagcca aaacaagaaa gaaqttatca gctcagacaa tctgcagtta lyw tds qnkk evissd nlql 541 cccgagetca ageagaagag ttctaataca agegetggge caactgtgce egacagagae pel k q k s s n t s a q p t v p d r d 601 aatgatggaa teeetgatag tetagaggtt gagggataca eggtagatgt caagaacaaa ndg ipd slevegy tvd vknk 661 aggacttttc tetegeettg gatetcaaat atecatgaga agaagggget taccaagtae rtflsp wisnihe kkg ltky aagteeteee eegagaagtg gtetaceget teegateeat atagegattt egagaaggte kss pek wstasd pysd fekv 781 acaggoogga tegataaaaa tgtgteteea gaggetagac accceetggt ageageetae tgridknvspearhplvaay 841 ccgattgtac acgtggacat ggagaacatc attctaagca aaaacgagga ccagtccaca piv hvd meni ils kne dqs t 901 caaaacactg actoogagac cogcaccata totaaaaaca ccagtactto aaggacccac q n t d s e tr t i s k n t s t s r t h 961 acctetgaag tgeacggeaa tgeggaagte catgeategt tittegatat tggtggetee tse vhg naevhas ffdiggs

1021 gtgtcagccg gctttagcaa tagcaactcc tcgacggttg ccattgacca ctcactgtca v s a g f s n s n s s t v a i d h s l s

1081 ttagcaggtg agaggacttg ggctgaaact atgggtctga ataccgccga tacggcccgg lagert waet mgl nta dtar 1141 ctcaacgcaa atattoggta cgtcaacaca gggactgctc ctatatataa cgtgctgcct lnanir yvnt gtapiy nvlp 1201 acgacaagtc ttgtcctggg caaaaatcag accctcgcaa ccattaaggc aaaggaaaat tts l v l g k n q t l a t i k a k e n 1261 cagetgagee agatectege cectaacaac tattatecat ceaaaaattt ageecceata qls qil apnn yyp skn lapi 1321 gccctgaacg cccaggacga cttttcctct acccccataa Ctatgaatta caatcagttc alnaqd dfsstpitmnyngf 1381 ctqqaqctqq aaaaqacqaa qcaqctqaqa ctaqacaccg atcaggtgta tqqaaacata lelekt kqlr ldt dqv yqn i 1441 gegacatata actttgagaa eggeegettg egegtegaca etgggteaaa etggtetgaa aty nfe ngrvrvd tgs nwse 1501 gttctgccgc aaattcaaga gacaaccgcc agaattatct ttaatgggaa ggacttgaac vlp qiq ettarii fn q k d l n 1561 cttgtegaac gtagaattge egeegtgaac eecagtgate cactegagac gactaaaccq lverriaavn psd ple ttkp 1621 gatatgacac tgaaagaggc tetgaagatt geetteggat teaacgaace taatggcaat dmt lke alki afg fne png n 1681 ttgcagtate aggggaaaga cateacagag tttgatttea atttegatea geagaettee lqy qqk dite fdf nfd qqts 1741 caaaatatca aaaatcagtt ggcagagctg aatgccacca atatctacac ggttctcgat qni knqlael nat niy tvld 1801 aaaatcaaac ttaacgccaa gatgaacata ttgattcgag acaaacgctt ccactacgac kik lna kmni lir dkr fhyd 1861 cqcaacaata taqccqtaqq cqctqatqaq tctqtcqtca aqqaqqctca taqqqaaqtt rnn i av qade svv kea hrev 1921 atcaacagca gtactgaagg gctgttactt aatatcgaca aggacattcg gaagatcctg ins steqlll nid kdi rkil 1981 teegggtata tegtggagat egaggatace gagggeetga aggaagteat taacgacege sgy ive iedt egl kev indr 2041 tatgatatgc tgaacatttc cagcttacga caggacggta agacatttat tgactttaaa ydm lnisslr qdg ktf idfk

2101 aagtataacg acaagctacc cctgtacatt tccaacccaa attacaaagt taatgtgtat k y n d k l p l y i s n p n y k v n v y

Figure 3C

- 2161 getgtaacca aggagaaca aatcatcaat ccaagcgaga acggcgatac cagcacaaat a v:t k e n t i i n p s e n g d t s t n
- 2221 ggaatcaaaa agatccttat atttagtaaa aaaggctacg agatcggttg aggatcc g i k k i l i f s k k g γ e i g -

- 1 gatategeca ecatggatge aatgaagaga gggetetget gtgtgetget getgttgtga
 m d a m k r q l c c v l l l c q
- 61 gcagtottog tttogoccag cgccggcggg catggggacg ttggcatgca tgtgaaagaa a v f v s p s a g g h g d v g m h v k e
- 121 aaggagaaaa acaaggacga aaacaagggt aaagacgaag aacgtaataa aacacaggag $k \ e \ k \ n \ k \ d \ e \ n \ k \ r \ k \ d \ e \ e \ r \ n \ k \ t \ g \ e$
- 181 gaacacttaa aggagatcat gaagcacata gtaaagattg aggtaaaagg cgaagaggct
 e h l k e i m k h i v k i e v k g e e a
- 241 gtaaagaagg aggcagcaga aaaactgttg gagaaggtgc cttctgacgt cttagagatg v k k e a a e k l l e k v p s d v l e m
- 361 ctcgaagctc tctccgagga caagaaaag attaaagaca tctacgggaa ggatgcctta
 l e a l s e d k k k i k d i y g k d a l
- ttgcacgagc actacgttta cgcaaaggag ggctatgagc ccgtgctcgt tattcagagt
 l h e h y v y a k e g y e p v l v i q s
- 481 agtgaggact acgtcgagaa taccgagaaa gctctgaatg tgtattacga gatcggaaag s e d y v e n t e k a l n v y y y e i g k
- 541 attetgteec gggacateet gteeaaaate aaceageeat aceagaaatt eettgatgtt
 i l s r d i l s k i n q p y q k f l d v
- 601 cttaacacaa tcaaaaacgc gtcagatagc gacgggcagg atcttctgtt tacaaatcaa
 l n t i k n a s d s d q q d l l f t n g
- 721 caggaggtgt togccaaggc atttgcgtac tatatcgaac cocagcatcg cgatgtgctc q e v f a k a f a y y i e p q h r d v l
- 781 cagctgtacg ccccggaggc atttaactac atggacaaat tcaatgaaca ggagattaat q 1 y a p e a f n y m d k f n e q e i n
- 841 ctgtctctgg aggaactgaa agaccagagg atgctctccc ggtatgaaaa gtgggaaaag l s l e e l k d q r m l s r y e k w e k
- 961 aaaaagttgc agattcccat cgagcctaag aaagatgata taatacactc tctaagccag
 k k l q i p i e p k k d d i i h s l s q
- 1021 gaggagaagg aactcctgaa gcggatacaa atcgactcat ccgatttcct tagcacagaa
 e e e k e l l k r i q i d s s d f l s t e

1081 qaqaaqqaqt ttctaaaaaa acttcaqata qatattaqaq attcactqaq cqaqqaaqaq eke flk klqi dir dsl seee 1141 aaggagetge teaacegaat teaagtegat agttegaace cettgteaga aaaagagaag kellnriqvd ssn pls 1201 gaattootga aaaagttgaa gotogacato cagoogtacg atattaatca goggotacaa efl k k l k l d i q p y d i n q r l q 1261 gacaccggcg gtctgattga tagccccagc atcaaccttg acgtacggaa gcaatataag dtg gli dsps inl dvr k qyk 1321 egggacatte aaaatatega egecetatta cateaateca taggeteeac getatacaat rdiqnidall hqsigstlyn 1381 aaaatotato tatacgaaaa catgaatatt aacaatotca cogotacact gggagoggac kiy lyenmni nnl tat lgad 1441 ctggtcgata gtacagacaa cacaaagata aacagaggta ttttcaacga attcaaaaag l v d s t d n t k i n r g i f n e f k k 1501 aactttaagt attegateag cagtaactat atgattgttg acatcaatga acggcccgca nfkysissny miv din erpa 1561 ttaqacaatq agaqqttgaa qtqqaqaatt caactqagtc ctqatactag qgccqqctat ldnerlkwriglspdtragy 1621 ctqqaqaacq qqaaactqat cttacaqcqa aacatcqggc tggagatcaa ggatgtqcag lengklilgrnigleikdvg 1681 attateaage agagegaaaa agaatacatt egeategaeg eeaaggtggt geetaagtea iik qse keyi ridak v v p k s 1741 aagategata eeaagateea ggaageteag eteaacatta accaggagtg gaataaaget ki-d t ki qeaq l n i n qe w n k a 1801 cttggtctgc caaaatacac caaacttatc acctttaatg tgcacaacag gtatgcctct lgl pky tkli tfn vhn ryas 1861 aatatogtog agtoagoata cotgattoto aatgaatgga agaacaatat toagtotgac nivesaylil new knn igsd 1921 ctgatcaaga aggtcacgaa ttatctggtg gacggaaatg gcagattcgt gttcaccgac lik kvt nylv dgn grf vftd 1981 ataactttqc caaacattqc cqaqcaatac actcatcaqq atqaaattta cqaqcaaqtc

- j

h s k g l y v p e s r s i l l h g p s k

2101 ggggttgagc ttcgaaacga ttctgaggga tttatcgctg actttggagc cgctgtggat
g v e l r n d s e g f i a d f g a a v d

i t l p n i a e q y t h q d e i y e q v q v q d e i y e q v q v q d catccaaag gtctgtatgt tccaqagtca aqatcqattc tqctccatqq tccatccaaa

Figure 4C

- 2161 gactacgccg gatacctgtt ggataagaat cagtctgate tegtgacaaa tagcaaaaaa d y a g y l l d k n q s d l v t n s k k
- 2221 ttcatagata ttttcaagga ggaaggagt aacctgactt cctatggccg cacgaacgag f i d i f k e e g s n l t s y g r t n e
- 2281 gctgaatttt ttgcggaagc ctttagactt atgcacagca ccgaccatgc tgaaaggttg
 a e f f a e a f r l m h s t d h a e r l
- 2341 aaggtgcaaa agaatgcccc taaaaccttc cagttcataa atgaccagat caagttcatc k v q k n a p k t f q f i n d q i k f i
- 2401 atcaactctt gaggatcc i n s -

- 1 gatatogoca coatggatgo aatgaagaga gggototgot gtgtgotgot gotgtgtgga m d a m k r g l c c v l l l c g
- 61 gcagtetteg tttegeccag egeeggeggg catggggaeg ttggcatgca tgtgaaagaa a v f v s p s a g g h g d v g m h v k e
- 121 aaggagaaaa acaaggacga aaacaagggt aaagacgaag aacgtaataa aacacaggag k e k n k d e n k r k d e e r n k t g e
- 181 gaacacttaa aggagatcat gaagcacata gtaaagattg aggtaaaagg cgaagaggct e h l k e i m k h i v k i e v k g e e a
- 241 gtaaagaagg aggcagcaga aaaactgttg gagaaggtgc cttctgacgt cttagagatg v k k e a a e k l l e k v p s d v l e m
- 301 tataaggcca teggeggtaa gatetatate gtggaeggag acateactaa acacatatet y k a i g g k i y i v d g d i t k h i s
- 361 ctcgaagctc tctccgagga caagaaaaag attaaagaca tctacgggaa ggatgcctta
 l e a l s e d k k k i k d i y q k d a l
- 421 ttgcacgage actacgttta cgcaaaggag ggctatgage ccgtgctcgt tattcagagt
 l h e h y v y a k e g y e p v l v i g s
- 481 agtgaggact acgtcgagaa taccgagaaa gctctgaatg tgtattacga gatcggaaag s e d y v e n t e k a l n v y y e i q k
- 541 attetgtee gggacatect gteeaaate aaccagecat accagaaatt cettgatgtt
 i l s r d i l s k i n q p y q k f l d v
- 601 cttaacacaa tcaaaaacgc gtcagatagc gacgggcagg atcttctgtt tacaaatcaa
 1 n t i k n a s d s d g g d l l f t n g
- 661 ctcaaggaac accccactga tttcagcgtg gagttcctcg agcagaattc taacgaagtc $1 \ k \ e \ h \ p \ t \ d \ f \ s \ v \ e \ f \ 1 \ e \ q \ n \ s \ n \ e \ v$
- 721 caggaggtgt tcgccaaggc atttgcgtac tatatcgaac cccagcatcg cgatgtgctc q e v f a k a f a y y i e p q h r d v l
- 781 cagctgtacg ccccggaggc atttaactac atggacaaat tcaatgaaca ggagattaat q l y a p e a f n y m d k f n e q e i n
- 841 ctgtctctgg aggaactgaa agaccagagg atgctctccc ggtatgaaaa gtgggaaaag $1\ s\ 1\ e\ e\ 1\ k\ d\ q\ r\ m\ 1\ s\ r\ y\ e\ k\ w\ e\ k$
- 901 atcaaacagc attaccagca ttggtccgac tccctgtcag aagaggggcg cggcctgttg
 i k q h y q h w s d s l s e e g r g l l
 - 961 aaaaagttgc agattcccat cgagcctaag aaagatgata taatacactc tctaagccag k k l q i p i e p k k d d i i h s l s q
 - 1021 gaggagaagg aactootgaa goggatacaa atogactoat cogatttoot tagcacagaa e e k e l l k r i g i d s s d f l s t e

- 1081 gagaaggagt ttctaaaaaa acttcagata gatattagag attcactgag cgaggaagag e k e f l k k l g i d i r d s l s e e e
- 1141 aaggagetge teaacegaat teaagtegat agttegaace cettgteaga aaaagagaag k e l l n r i q v d s s s n p l s e k e k
- 1201 gaatteetga aaaagttgaa getegacate eageegtaeg atattaatea geggetaeaa e f l k k l k l d i g p y d i n g r l g
- 1261 gacaccggcg gtctgattga tagccccagc atcaaccttg acgtacggaa gcaatataag
 d t g g l i d s p s i n l d v r k q y k
- 1321 cgcgacattc aaaatatcga cgccctatta catcaatcca taggctccac gctatacaat r d i q n i d a l l h q s i g s t l y n
- 1381 aaaatctatc tatacgaaaa catgaatatt aacaatctca ccgctacact gggagcggac k i y l y e n m n i n n l t a t l q a d
- 1441 ctggtcgata gtacagacaa cacaaagata aacagaggta ttttcaacga attcaaaaag l v d s t d n t k i n r g i f n e f k k
- 1501 aactttaagt attegateag dagtaactat atgattgttg acatcaatga acggceegca n f k v s i s s n y m i v d i n e r p a
- 1621 ctggagaacg ggaaactgat cttacagcga aacatcgggc tggagatcaa ggatgtgcag l e n g k l i l q r n i g l e i k d v q
- 1681 attatcaagc agagcgaaaa agaatacatt cgcatcgacg ccaaggtggt gtagggatcc i i k g s e k e v i r i d a k v v -

Figure 6

- 1 gatategeca ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga m d a m k r g l c e v l l l c g
- 61 gcagtetteg tttegeceag egeeggeggg catggggaeg ttggcatgca tgtgaaagaa a v f v s p s a g g h g d v g m h v k e
- 121 aaggagaaaa acaaggacga aaacaagcgt aaagacgaag aacgtaataa aacacaggag k e k n k d e n k r k d e e r n k t q e
- 181 gaacacttaa aggagatcat gaagcacata gtaaagattg aggtaaaagg cgaagaggct e h l k e i m k h i v k i e v k q e e a
- 241 gtaaagaagg aggcagcaga aaaactgttg gagaaggtgc cttctgacgt cttagagatg v k k e a a e k l l e k v p s d v l e m
- 301 tataaggcca teggeggtaa gatetatate gtggaeggag acateactaa acacatatet
 y k a i q q k i y i v d q d i t k h i s
- 361 ctcgaagctc tctccgagga caagaaaaag attaaagaca tctacgggaa ggatgcctta l e a l s e d k k k i k d i y q k d a l
- 421 ttgcacgagc actacgttta cgcaaaggag ggctatgagc ccgtgctcgt tattcagagt
 l h e h y v y a k e q y e p v l v i g s
- 481 agtgaggact acgtcgagaa taccgagaaa gctctgaatg tgtattacga gatcggaaag s e d y v e n t e k a l n v y y e i g k
- 541 attetgteee gggacateet gteeaaaate aaceageeat aceagaaatt eettgatgtt
 i l s r d i l s k i n q p y q k f l d v
- 601 cttaacacaa tcaaaaacgc gtcagatagc gacgggcagg atcttctgtt tacaaatcaa l n t i k n a s d s d g q d l l f t n q
- ctcaaggaac accccactga tttcagcgtg gagttcctcg agcagaattc taacgaagtc l k e h p t d f s v e f l e q n s n e v
- 721 Caggaggtgt tegecaagge attttgagga tee

- 1 gatategeca ecatggatge aatgaagaga gggetetget gtgtgetget getgttgtgga md amkr qlc cvl llc q 61 gcaqtcttcq tttcgcccag cagcgctggg ccaactgtgc ccgacagaga caatgatgga avf vsp ssag ptv pdr dndg 121 atccctgata gtctagaggt tgagggatac acggtagatg tcaagaacaa aaggactttt ipd sle vegy tvd vkn krtf 181 ctctcqcctt qqatctcaaa tatccatqaq aaqaaqqqqc ttaccaaqta caaqtcctcc lsp wis nihe kkg ltk ykss 241 cccgagaagt ggtctaccgc ttccgatcca tatagcgatt tcgagaaggt cacaggccgg pek wstasdpysd fek vtgr 301 ategataaac aggtgtetee agaggetaga caccecetgg tageageeta ceegattgta idk qvs pear hpl vaa ypiv cacgtggaca tggagaacat cattotaago aaaaacgagg accagtocac acaaaacact hvd men iils kne dqs tqn t 421 gacteegaga eeegeaceat atetaaacag accagtactt caaggaceca cacetetgaa d s e trt i s k q t s t s r t h t s e 481 gtgcacggca atgcggaagt ccatgcatcg tttttcgata ttggtggctc cgtgtcagcc vhg nae vhas ffdigg svsa 541 ggetttagea atagecagte etegaeggtt gecattgace acteaetgte attageaggt qfs nsq sstv aid hsl slag 601 gagaggactt gggctgaaac tatgggtctg aatacegeeg atacggeeeg gctcaaegca ert wae tmgl ntadtarlna aatatteggt aegteaacae agggaetget cetatatata aegtgetgee taegacaagt nir yvn tgtapiy nvl ptts 721 ettgteetgg geaaacagea gaccetegea accattaagg caaaggaaaa teagetgage lvlgkqqtlatikakenqls
- 781 cagatectog.cocotaacaa etattateca tecaaaaaatt tagececcat agecetgaac q i l a p n n y y p s k n l a p i a l n 841 geceaggacg acttitecte tacceccata actatgaatt acaatcagt cetgagetg a q d d f s s t p i t m n y n q f l e l
- 901 gaaaagacga agcagctgag actagacacc gatcaggtgt atggaaacat agcgacatat e k t k q l r l d t d q v y g n i a t y
- 961 aactttgaga acggeegegt gegegtegae actgggteae agtggtetga agttetgeeg n f e n g r v r v d t g s q w s e v l p
- 1021 caaattcaag agacaaccgc cagaattatc tttaatggga aggacttgaa ccttgtcgaa q i q e t t a r i i f n g k d l n l v e

- 1081 cgtagaattg ccgccgtgca gcccagtgat ccactcgaga cgactaaacc ggatatgaca r r i a a v q p s d p l e t t k p d m t
- 1141 ctgaaagagg ctctgaagat tgocttcgga ttcaacgaac ctaatggcaa tttgcagtat $1 \quad k \quad e \quad a \quad 1 \quad k \quad i \quad a \quad f \quad g \quad f \quad n \quad e \quad p \quad n \quad g \quad n \quad 1 \quad q \quad y$
- 1201 caggggaaag acatcacaga gtttgattte aatttegate agcagactte ccaaaatate $q\ q\ k\ d\ i\ t\ e\ f\ d\ f\ n\ f\ d\ q\ q\ t\ s\ q\ n\ i$
- 1261 aaaaatcagt tggcagagct gcaggccacc aatatctaca cggttctcga taaaatcaaa k n q l a e l q a t n i y t v l d k i k
- 1321 ettaaegeea agatgaacat attgattega gacaaaeget teeactaega eegeaacaat $1 \ n \ a \ k \ m \ n \ i \ 1 \ i \ r \ d \ k \ r \ f \ h \ y \ d \ r \ n \ n$
- limit at aggregate groups and groups groups at a suggraph at a suggraph and groups and groups and groups at a suggraph at a sug
- 1441 agtactgaag ggctgttact taatatcgac aaggacatte ggaagatect gteegggtat s t e g l l l l n i d k d i r k i l s g y
- 1501 atcgtggaga tcgaggatac cgagggcctg aaggaagtca ttaacgaccg ctatgatatg i v e i e d t e g l k e v i n d r y d m
- 1621 gacaagetae eeetgtacat ticcaaceca aattacaaag ttaatgtgta tgetgtaace d k l p l y i s n p n y k v n v y a v t
- 1681 aaggagaaca caatcatcca gccaaggag aacggcgata ccagcacaaa tggaatcaaa k e n t i i q p s e n g d t s t n g i k
- 1741 aagateetta tatttagtaa aaaaggetae gagateggtt gaggatee k i l i f s k k g y e i g -

1 gatategeea eeatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga mdamkr qlc cvl llc q · 61 gcagtetteg tttegeccag egeeggeggg catggggaeg ttggcatgea tgtgaaagaa avf vsp sagg h q d v q m 121 aaggagaaaa acaaggacga aaacaagcgt aaagacgaag aacgtcagaa aacacaggag kek nkdenkr kde erg kt qe 181 gaacacttaa aggagatcat gaagcacata gtaaagattg aggtaaaagg cgaagaggct ehl keim khi v ki ev k qee a 241 gtaaagaagg aggcagcaga aaaactgttg gagaaggtgc cttctgacgt cttagagatg vkk eaa ekll ekv psd vle m 301 tataaggeea teggeggtaa gatetatate gtggaeggag acateactaa acacatatet yka igg kiyi vdg dit 361 ctcgaagete tetecgagga caagaaaaag attaaagaca tetacgggaa ggatgeetta leal se dkkk ikd i yg kdal 421 ttgcacqaqc actacqttta cgcaaaqqaq qgctatqaqc ccqtqctcqt tattcaqaqt lhehyvyake gyepvl viqs 481 aqtqaqqact acqtcqaqaa taccqaqaaa qctctqaatq tqtattacqa qatcqqaaaq sed y ve ntek aln v y y e i g k 541 attetgteee gggacateet gteeaaaate aaceageeat aceagaaatt cettgatgtt ils rdilski nqpyqk fldv 601 cttaacacaa tcaaacaggc gtcagatagc gacgggcagg atcttctgtt tacaaatcaa 1 ntikqasds dgqdll ftng 661 ctcaaggaac accccactga tttcagcgtg gagttcctcg agcagaattc taacgaagtc lke hpt dfsvefleqn 721 caggaggtgt tegecaagge atttgegtac tatategaac eccageateg egatgtgete qev fakafay yie pqh rdv 1 781 cagetgtacg ecceggagge atttaactac atggacaaat teaatgaaca ggagatteag qly ape afny m dk fne qeiq 841 ctgtctctgg aggaactgaa agaccagagg atgctctccc ggtatgaaaa gtgggaaaag l s l e e l k d q r m l s r y e k w e k 901 atcaaacagc attaccagca ttggtccgac tccctgtcag aagaggggcg cggcctgttg ikq hyqhwsd sls eeg rgll 961 aaaaagttgc agattcccat cgagcctaag aaagatgata taatacactc tctaaqccag kkl qipiepk kdd iih slsq

1021 gaggagaagg aactcctgaa gcggatacaa atcgactcat ccgattcct tagcacagaa e e e k e l l k r i q i d s s d f l s t e

1081 qaqaaqqaqt ttctaaaaaa acttcagata gatattagaq attcactqaq cqaqqaaqaq eke flk klqi dir dsl seee 1141 aaggagetge teaacegaat teaagtegat agttegaace cettgteaga aaaagagaag kellnriqvd ssn pls ekek 1201 gaattootga aaaagttgaa gotogacato cagoogtacg atattaatca goggotacaa eflkkl kldiqpy din qrlq 1261 gacaceggeg gtetgattga tagececage ateaacettg acgtacggaa gcaatataag dtq qli dsps inl dvr k qyk 1321 egegacatte aaaatatega egeeetatta cateaateea taggeteeae getatacaat rdiqnidall hqsigstlyn 1381 aaaatotato tatacgaaaa catgaatatt aaccagotca cogotacact gggagoggac kiy lyenmni nql tatlqad 1441 ctqqtcqata qtacaqacaa cacaaagata aacagaggta ttttcaacga attcaaaaag lvd std ntki nrg i f n. e f k k 1501 aactttaagt attcgatcag cagtaactat atgattgttg acatcaatga acggcccgca nfk ysissny miv din erpa 1561 ttaqacaatq aqaqqttgaa gtggagaatt caactgagtc ctgatactag ggccggctat ldnerlkwriglspdtragy 1621 ctqqaqaacq qqaaactqat cttacagcga aacatcgggc tggagatcaa ggatgtgcag lengklilqrnigleikdvq 1681 attatcaagc agagcgaaaa agaatacatt cgcatcgacg ccaaggtggt gcctaagtca iik qse keyi ridak v v p k s 1741 aaqatcqata ccaaqatcca qqaaqctcaq ctcaacatta accaqqaqtq qaataaaqct kid tki qeaq lni nqe wnka cttggtctgc caaaatacac caaacttatc acctttaatg tgcacaacag gtatgcctct lqlpkytklitfnvhnryas 1861 aatategteg agteageata eetgattete aatgaatgga agaacaatat teagtetgae nivesaylil new knn iqsd 1921 otgatcaaga aggtcacgaa ttatotggtg gacggaaatg gcagattegt gttcaccgac lik kvt nylv dgn grf vftd 1981 ataactttgc caaacattgc cgagcaatac actcatcagg atgaaattta cgagcaagtc itl p n i a e q y t h q d e i yeqv 2041 cactocaaag gtetgtatgt tecagagtea agategatte tgetecatgg tecatecaaa h s k gly vpes rsi llh gpsk 2101 ggggttgagc ttcgacagga ttctgaggga tttatcgctg actttggagc cgctgtggat gvelrqdseqfiadfgaavd

1

Figure 8C

- 2161 gactacgccg gatacctgtt ggataagcag cagtctgate tegtgacaaa tagcaaaaaa d y a g y 1 1 d k q q s d 1 v t n s k k
- 2221 ttcatagata ttttcaagga ggaagggagt cagctgactt cctatggccg cacgaacgag f i d i f k e e g s q l t s y g r t n e
- 2281 gctgaatttt ttgcggaagc ctttagactt atgcacagca ccgaccatgc tgaaaggttg a e f f a e a f r 1 m h s t d h a e r 1
- 2341 aaggtgcaaa agaatgcccc taaaaccttc cagttcataa atgaccagat caagttcatc k v q k n a p k t f q f i n d q i k f i
- 2401 atcaactctt gaggatcc

ins -

Figure 9A

. .

•		
TPA-Human B.a PA	PA	agcgctgggccaactgtgcccgacagagacaatgatggaatccctgatagtctagagggttgagggatacatatgt.
TPA-Human B.a PA	PA	cggtagatgtcaagaacaaaaggacttttctctcgccttggatctcaaatatccatgagaagaaggggct,taattaattttttttttttttttttttttt
TPA-Human B.a PA	PA	taccaagtacaagtcctcccccgagaagtggtctaccgcttccgatccatatagcgatttcgagaaggtcaataattaaagcgtgct
TPA-Human B.a PA	PA	a cagge cgg at cgataaaaat gt gt ctc cag agg ctagacaccccct gg tag cagcctacccg at tgt acatgaatgttgtt
TPA-Human B.a PA	PA	acgtggacatggagaacatcattctaagcaaaaacgaggaCcagtccacacaaaacactgactccgagac.tatt.t.tctcatt.
TPA-Human B.a PA	PA	ccgcaccatatctaaaaacaccagtacttcaaggacccacacctctgaagtgcacggcaatgcggaagtcga.aaagtttcaagtattagatag
TPA-Human B.a PA	PA	$catgeategtttttegatattggtggeteegtgteageeggetttageaatageaaeteetegaeggttg\\ \dots \\ g.\dots.et\\ \dots .gagtata.a.\dots.t\\ \dots .tegtaagt.\dots.e.$
TPA-Human B.a PA	PA	$ccattgaccactcactgtcattagcaggtgagaggacttgggctgaaactatgggtctgaataccgccga\\ a. \dots.tt.\dots atc.\dotsgaa. \dots .t.a. \dots .t.a. \dots .t.a. \dots .t.a.$
TPA-Human B.a PA	PA	tacggcccggctcaacgcaaatattcggtacgtcaacacagggactgctcctatatata
TPA-Human B.a PA	PA	acgacaagtcttgtcctgggcaaaatcagacctcgcaaccattaaggcaaaggaaaatcagctgagcc ttcgt.agt.aaaagaa
TPA-Human B.a PA	PA	agatectegecectaacaactattatecatecaaaaatttagececeatagecetgaacgeceaggacga.aatattttt
TPA-Human B.a PA	PA	$\tt cttttcctctacccccataactatgaattacaatcagttcctggagctggaaaagacgaagcagctgagatcagtt.at.at.aaaaa$
TPA-Human B.a PA	PA	ctagacaccgatcaggtgtatggaaacatagcgacatataactttgagaacggccgcgtgcgcgtcgacattgaaagtaactaaaa.ggt.
TPA-Human B.a PA	PA	$ctgggtcaaactggtctgaagttctgccgcaaattcaagagacaaccgccagaattatctttaatgggaa, acg., \dots.ag., \dots.gt.a., \dots.a., \dots.tac.tct., \dots.a.,\\$
TPA-Human B.a PA	PA	$ggacttgaaccttgtcgaacgtagaattgccgccgtgaaccccagtgatccactcgagacgactaaaccg\\ atatgaagc_gaggtttt.aa$
TPA-Human B.a PA	PA	gatatgacactgaaagaggctctgaagattgccttcggattcaacgaacctaatggcaatttgcagtatc t.aact.aatt
TPA-Human B.a PA	PA	aggggaaagacatcacagagtttgatttcaatttcgatcagcagacttcccaaaatatcaaaaatcagtt .aacattaaat
TPA-Human B.a PA	PA	ggcagagctgaatgccaccaatatctacacggttctcgataaaatcaaacttaacgccaagatgaacataagat.acatcat.at.at.at.at.at.at.at.at.at.a
TPA-Human B.a PA	PA	ttgättcgagacaaacgcttccactacgaccgcaacaatatagccgtaggcgctgatgagtctgtcgtca aaat.gg.gt.t.t.t.t.t.a.atcat.gggaat.
TPA~Human B.a PA	PA	aggaggctcatagggaagttatcaacagcagtactgaagggctgttacttaatatcgacaaggacattcgaa.tttcgtca.a.g.ata.a.gt.at.t.ttt.aa.
TPA-Human B.a PA	PA	$\label{eq:gagatcctg} gaagatcctgtccgggtatatcgtggggatcgaggatccgc\\ aaat.aa.ttagtagtatagtatagtatagtagtatagtatagtatagtagtatagtagtatagtagtagtatagtatagtatagtatagtatagtatagtatagtatatagtatatata$
TPA-Human B.a PA	PA	$tatgatatgctgaacatttccagcttacgacaggacggtaagacatttattgactttaaaaagtataacg\\ \dots \dots$

' Figure 9B

TPA-Human PA	1622 acaagctacccctgtacatttccaacccaaattacaaagttaatgtgtatgctgtaaccaaggagaacac
B.a PA	2135 .tatgt.ataagttctgaattaa
TPA-Human PA	1692 aatcatcaatccaagcgagaacggcgataccagcacaaatggaatcaaaaagatccttatatttagtaaa
B.a PA	2205 tttttgttccggatt.actc
TPA-Human PA	1762 aaaggctacgagatcggttga
B.a PA	2275taa.a.

Figure 10A	
TPA-human fu	82 gaagtgaagcaagaaaatcgacttetgaaegagagegaaagtteateacagggtettete
B.a PA	88t.a.g.g.g.c.gt.at.a.t.ateateaagt.cgt.a.a
TPA-human fu	142 ggatactacttcagtgacttgaatttccaagcaccaatggtggtgactagtagcaccacc
B.a PA	148tttttctctc.tct.ta
TPA-human fu	202 ggcgatttgagcattcccagctctgagttggagaacattcccagcgaaaatcagtacttc
B.a PA	208 .gatctt.ta.a.a.tatcgc.a.tt
TPA-human fu	262 cagtctgctatctggtccggattcattaaggttaaaaagtccgacgaatatacatttgct
B.a PA	268 .atatc.agagt.t
TPA-human fu	322 acctcggcggataaccatgtgacaatgtgggtggacgaccaggaagtgatcaacaaggct
B.a PA	328t.cttaatat.t.t.a
TPA-human fu	382 tcaaactctaataaaatccggctcgagaaggggaggctctaccagatcaaaattcagtac
B.a PA	388ttca.at.aaaaat.at.aaat
TPA-human fu	442 cagcgggaaaaccctacagaaaaaggactcgatttcaagctgtactggacagatagccaa
B.a PA	448 .a.attt.gttctct
TPA-human fu	502 aacaagaagagttatcagctcagacaatctgcagttacccgagctcaagcagaagagt
B.a PA	508 .tagttctagttct.aaga.ta.a.a.a.a.a.a.
TPA-human fu	562 tctaatacaagcgctgggccaactgtgcccgacagagacaat
B.a.~ PA	568 .gctcaagaaaaagcgaagtatgtac.t
TPA-human fu	604 gatggaatccctgatagtctagaggttgagggatacacggtagatgtcaagaacaaaagg
B.a PA	628tcata.atta.ta
TPA-human fu B.a PA	664 acttttctctcgccttggatctcaaatatccatgagaagaaggggcttaccaagtacaag 688tt.a.at.a.ata
TPA-human fu	724 tectececegagaagtggtetacegetteegatecatatagegatttegagaaggteaca
B.a PA	748attaaagcgtgc.t.tat
TPA-human fu	784 ggccggatcgataaaaatgtgtotccagaggctagacacccctggtagcagcctacccg
B.a PA	808 .atga.a.aat.gt.gt.t
TPA-human fu	844 attgtacacgtggacatggagaacatcattctaagcaaaaacgaggaccagtccacacaa
B.a PA	868tatt.tctcattag
TPA-human fu	904 aacactgactccgagacccgcacatatctaaaaacaccagtacttcaaggacccacacc
B.a PA	928ttagtaga.aaagtttcaagtatt
TPA-human fu	964 totgaagtgcacggcaatgcggaagtocatgcatcgtttttcgatattggtggctccgtg
B.a PA	988 agataaggct
TPA-human fu B.a PA	$1024\ t cag ccg g ctttag caatag caact cctcg acgg ttg ccattg accact cactg t catta \\ 1048\taattcg., taagtcat., tatc$
TPA-human fu B.a PA	1084 gcaggtgagaggacttgggctgaaactatgggtctgaataccgccgataccgcccggctc 1108ga.a.a.a.a.a.a.a.a.a.a.a.a.a.a.a.
TPA-human fu B.a PA	$1144\ aacgcaaatattcggtacgtcaacacagggactgctcctatatata$
TPA-human fu	1204 acaagtettgteetgggcaaaaateagaeeetegcaaceattaaggcaaaggaaaateag
B.a PA	1228ttegt.agt.aaaagaatca

1324 ctgaacgcccaggacgacttttcctctacccccataactatgaattacaatcagttcctg

1348 t.a..t..a..a...t..cagt.....t..a..t..a..t..a..t..a..t..t

TPA-human fu

TPA-human fu

B.a.- PA

B.a.- PA

Figure 10B	
TPA-human fu	1384 gagctggaaaagacgaagcagctgagactagacaccgatcaggtgtatggaaacatagcg
B.a PA	1408t.aaaat.attga.agta
TPA-human fu	1444 acatataactttgagaacggccgcgtggcgtcgacactgggtcaaactggtctgaagtt
B.a PA	1468c.ta.taa.aaggt.acgagg
TPA-human fu	1504 ctgccgcaaattcaagagacaaccgccagaattatctttaatgggaaggacttgaacctt
B.a PA	1528 t.aat.at.g
TPA-human fu	1564 gtcgaacgtagaattgccgccgtgaaccccagtgatccactcgagacgactaaaccggat
B.a PA	1588 .a.a.gc.g.a.g.g.t.t.t.
TPA-human fu	1624 atgacactgaaagaggctctgaagattgccttcggattcaacgaacctaatggcaatttg
B.a PA	1648t.aa.cta.cttga.ca
TPA-human fu	1684 cagtatcaggggaaagacatcacagagtttgatttcaatttcgatcagcagacttcccaa
B.a PA	1708 .aaacat
TPA-human fu	1744 aatatcaaaaatcagttggcagagctgaatgccaccaatatctacacggttctcgataaa
B.a PA	1768gagat.acatcattat.a
TPA-human fu	1804 atcaaacttaacgccaagatgaacatattgattcgagacaaacgcttccactacgaccgc
B.a PA	1828t.atatttta.a
TPA-human fu	1864 aacaatatagccgtaggcgctgatgagtctgtcgtcaaggaggctcatagggaagttatc
B.a PA	1888 .t.ca.t.g.g.ga.a.t.
TPA-human fu	1924 aacagcagtactgaagggctgttacttaatatcgacaaggacattcggaagatcctgtcc
B.a PA	1948ttcgtcaagat.agt.atttaa.aaat.aa

2008 ..t....t..a..a..t..a...t..a..g..t..a...t..a..t...a.a...

2044 gatatgctgaacatttccagcttacqacaqqacqqtaaqacatttattqactttaaaaaq

2104 tataacgacaagctacccctgtacatttccaacccaaattacaaagttaatgtgtatgct

2164 gtaaccaaggagaacacaatcatcaatccaagcgagaacggcgataccagcacaaatgga

2188 ..t..t..a..a....t..t..t...t..t...t..g....t..t..c..c..g

2224 atcaaaaagatccttatatttagtaaaaaaggctacgagatcggttgaggatcc

2248g..a..tt.a..c...tc......t....a..a.a.a.-----

TPA-human fu

TPA-human fu

TPA-human fu

TPA-human fu

TPA-human fu

B.a.- PA

B.a.- PA

B.a.- PA

B.a.- PA

B.a.- PA

Figure 11A TPA-Human LF 82 gccqqcqqqcatqqqqacqttqqcatqcatqtqaaaqaaaaqqaqaaaaacaaqqacqaa Ba - LF 100 ..g....t...t..t..a..t....c..a.....g..a.....t..a..t...g TPA-Human LF 142 aacaagcqtaaaqacqaagaacgtaataaaacacaggaggaacacttaaaggagatcatq B.a.- LF TPA-Human LF 202 aagcacataqtaaaqattqaqqtaaaaqqcqaaqaqqctqtaaaqaaqqaqqcaqcaqaa B.a. - T.F 220 ..a. ...t...a.a.a.a.a...g.,g.,g.,a....t.,a.,a.,.... 262 aaactgttggagaaggtgccttctgacgtcttagagatgtataaggccatcggcggtaag TPA-Human LF B.a.- LF TPA-Human LF 322 atctatatcqtqqacqqaqacatcactaaacacatatctctcqaaqctctctccqaqqac Ba- LF 340 ..a....t...t..t..t..t..a.....t.a.....t.a....at.a..t..a..t TPA-Human LF 382 aagaaaaagattaaagacatctacgggaaggatgccttattgcacgagcactacgtttac B.a.- LF TPA-Human LF 442 gcaaaggagggctatgagcccgtgctcgttattcagagtagtgaggactacgtcgagaat B.a.- LF 460a.a.a.a...a...a..t..a..c..atc.tcq..a..t..t..a..a... TPA-Human LF 502 accgagaaagctctgaatgtgtattacgagatcggaaagattctgtcccgggacatcctg Ba~ LF 520 ..t.a.q.a....c..t....t.a.a..t....at.a..aa...t..tt.a TPA-Human LF 562 tocaaaatcaaccagccataccagaaattccttgatgttcttaacacaatcaaaacgcg B. a. - LF 580 agt....t..t..a.....t......tt.a....at.a..t..c..t....t..a TPA-Human LF 622 tcagatagcgacgggcaggatcttctgtttacaaatcaactcaaggaacaccccactgat B.a. - LF 640 ..t...tca..t..a..a.....t.a....t.....g..t.....t.....t.....a...c TPA-Human LF 682 ttcaqcqtqqaqttcctcqaqcagaattctaacqaagtccaqqaggtgttcqccaaqqca B.a.- LF 700 ..ttct..a..a...t.g..a..a...agc..t..g..a..a..a..a..t..g..a..t TPA-Human LF 742 tttgcgtactatatcgaaccccagcatcgcgatgtgctccagctgtacgccccggaggca B.a. - LF 760a.t.....q.a.....t....tt.a....t.t.a....t.a....t.a...t TPA-Human LF 802 tttaactacatggacaaattcaatgaacaggagattaatctgtctctggaggaactgaaa B.a. - LF 820t.....t...t..c....a..a..a..a..ct....a...t... TPA-Human LF 862 gaccagaggatgctctcccggtatgaaaagtgggaaaagatcaaacagcattaccagcat B.a.- LF 880 ..t..ac.....q..aa.a......a.....a......a.....c..t..a..c TPA-Human LF 922 tggtccgactccctgtcagaagaggggcgcggcctgttgaaaaagttgcagattcccatc B.a. - LF 940 ...ag...t..tt.a..t....a..aa.a..a..t..a.....c......t..t TPA-Human LF 982 gagcctaagaaagatgatataatacactctctaagccaggaggagaaggaactcctgaag B.a. - LF TPA-Human LF 1042 cggatacaaatcgactcatccgatttccttagcacagaagagaaggagtttctaaaaaaa Ba-LF TPA-Human LF B.a - LF 1120 ..a. a. t. c. t. ... tt.atct..a. ... a. .a. ... tt.a. .ta...a TPA-Human LF 1162 caaqtcqataqttcqaaccccttqtcaqaaaaaqaqaaqqaattcctqaaaaaqttqaaq Ba-LF 1180 ..g..g.....agt..t..t..a..t......a..a..g..tt.a.....c...a TPA-Human LF B.a.- LF 1240 ..t..t..t..a..a..t............aa..t.g.....t..a..a..gt.a..... TPA-Human LF 1282 agccccagcatcaaccttqacqtacqqaaqcaatataaqcqcqacattcaaaatatcqac

1360 ..tt.....t..aaqt..ct.q......t...t.q..t....t

B.a.- LF

B.a. - LF

TPA-Human LF

Figure 11B	
TPA-Human LF	1402 atgaatattaacaatctcaccgctacactgggagcggacctggtcgatagtacagacaac
B.a LF	1420ct.ctaacattt.attcctt
TPA-Human LF	1462 acaaagataaacagaggtattttcaacgaattcaaaaagaactttaagtattcgatcagc
B.a LF	1480t.attttat.caagtttct
TPA-Human LF	1522 agtaactatatgattgttgacatcaatgaacggcccgcattagacaatgagaggttgaag
B.a LF	1540tc.ta
TPA-Human LF	1582 tggagaattcaactgagtcctgatactagggccggctatctggagaacgggaaactgatc
B.a LF	1600ct.atcaac.aaat.aa.tagta
TPA-Human LF	1642 ttacagcgaaacatcgggctggagatcaaggatgtgcagattatcaagcagagcgaaaaa
B.a LF	1660aata.a.aa.a.a.a.tatc
TPA-Human LF	1702 gaatacattegeategaegeeaaggtggtgeetaagteaagategataceaagateeag
B.a LF	1720t.aa.gtt.gaaaagtaaa.
TPA-Human LF	1762 gaageteageteaacattaaccaggagtggaataaagetettggtetgecaaaatacace
B.a LF	1780at.at.ataat.a.gt.agt.ata
TPA-Human LF	1822 aaacttatcacctttaatgtgcacaacaggtatgcctctaatatcgtcgagtcagcatac
B.a LF	1840 .gt.ac.c.ct.t.aa.ct.a.aagtt.t
TPA-Human LF	1882 ctgattctcaatgaatggaagaacaatattcagtctgacctgatcaagaaggtcacgaat
B.a LF	1900 t.aat.ga.t.taagt.t.aa.aaa
TPA-Human LF B.a LF	$194\overset{?}{2}\ tatetggtggaeggaaatggeagattegtgtteaeegaeataaetttgeeaaeattgee\\ 1960\et.at.,ttatttttt$
TPA-Human LF	2002 gagcaatacactcatcaggatgaaatttacgagcaagtccactccaaaggtctgtatgtt
B.a LF	2020atag.at.t.t.agt.agt.a
TPA-Human LF	2062 ccagagtcaagatcgattctgctccatggtccatccaaaggggttgagcttcgaaacgat
B.a LF	2080acc.tt.at.aatataat.aa.g.,t
TPA-Human LF	2122 tctgagggatttatcgctgactttggagccgctgtggatgactacgccggatacctgttg
B.a LF	2140 agtacacacatt.t.t.t.t
TPA-Human LF	2182 gataagaatcagtctgatctcgtgacaaatagcaaataattcatagatattttcaaggag
B.a LF	2200c.at.attcttta

2242 gaagggagtaacctgacttcctatggccgcacgaacgaggctgaattttttgcggaagcc

2260tt.a....g....ga.a..a..t..a..g.....a...

2302 tttagacttatgcacagcaccgaccatgctgaaaggttgaaggtgcaaaagaatgccct

2320gt.a.....ttct..g.........c.t..a..a..t.....a...t...g

 $2362\ aaaaccttccagttcataaatgaccagatcaagttcatcatcaactcttgaggatcc$

2380t...a..t..t..c..t...t...t..t....a.a.-----

TPA-Human LF

TPA-Human LF

TPA-Human LF

B.a.- LF

B.a.- LF

B.a.- LF

Figure	12
--------	----

TPA-Human PA	1 mdamkrglccvll1cgavfvspssagptvpdrdndgipdslevegytvdvknkrtf1spw
Sugar minus	1
TPA-Human PA Sugar minus	61 isnihekkgltkyksspekwstasdpysdfekvtgridknyspearhplvaaypivhvdm 61
TPA-Human PA	121 eniilsknedgstqntdsetrtiskntstsrthtsevhgnaevhasffdiggsvsagfsn
Sugar minus	121qq
TPA-Human PA	181 snsstvaidhslslagertwaetmglntadtarlnaniryvntgtapiynvlpttslvlg
Sugar minus	181 . ${f q}$
TPA-Human PA	241 kngtlatikakenqlsqilapnnyypsknlapialnaqddfsstpitmnynqflelektk
Sugar minus	241 .q
TPA-Human PA	301 qlrldtdgvygniatynfengrvrvdtgsnwsevlpqiqettariifngkdlnlverria
Sugar minus	301qq
TPA-Human PA-	361 avnpsdplettkpdmtlkealkiafgfnepngnlqyqgkditefdfnfdqqtsqniknql
Sugar minus	361q
TPA-Human PA Sugar minus	421 aelnatniytvldkiklnakmnilirdkrfhydrnniavgadesvvkeahrevinssteg 421 qqq
TPA-Human PA	481 lllnidkdirkilsgyiveiedteglkevindrydmlnisslrqdgktfidfkkyndklp
Sugar minus	481qq
TPA-Human PA	541 lyisnpnykvnvyavtkentiinpsengdtstngikkilifskkgyeig
Sugar minus	541qq.

Figure 13

TPA-Human LF Sugar minus	${\it mdamkrg1ccvll1cgavfvspsagghgdvgmhvkekeknkdenkrkdeernktqeehlk} \\ \dots q. \dots \\ q. \dots \\ q. \dots$
TPA-Human LF Sugar minus	eimkhivkievkgeeavkkeaaekllekvpsdvlemykaiggkiyivdgditkhisleal
TPA-Human LF Sugar minus	sedkkkikdiygkdallhehyvyakegyepvlviqssedyventekalnvyyeigkilsr
TPA-Human LF Sugar minus	$\begin{tabular}{ll} dilsk inqpyqkfldvlntiknasdsdgqdllftnqlkehptdfsvefleqnsnevqevf \\ \\ q \\$
TPA-Human LF Sugar minús	akafayyiepqhrdvlqlyapeafnymdkfneqeinlsleelkdqrmlsryekwekikqh q
TPA-Human LF Sugar minus	yqhwsdslseegrgllkklqipiepkkddiihslsqeekellkriqidssdflsteekef
TPA-Human LF Sugar minus	$1 kkl qidirdsl see ekelln riqvds snpl se kekeflkklkldiq pydin qrlqdtgg \\ \\$
TPA-Human LF Sugar minus	lidspsinldvrkqykrdiqnidallhqsigstlynkiylyenmninnltatlgadlvds qq
TPA-Human LF Sugar minus	${\tt tdntkinrgifnefkknfkysissnymivdinerpaldnerlkwriqlspdtragyleng}$
TPA-Human LF Sugar minus	${\tt klilqrnigleikdvqiikqsekeyiridakvvpkskidtkiqeaqlninqewnkalglp}$
TPA-Human LF Sugar minus	kytklitfnvhnryasnivesaylilnewknniqsdlikkvtnylvdgngrfvftditlp
TPA-Human LF Sugar minus	$\label{eq:niaeqythqdeiyeqvhskglyvpesrsillhgpskgvelrndsegfiadfgaavddyag} \dots \dots q \dots q \dots$
TPA-Human LF Sugar minus	$\label{eq:control_problem} y \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\$
TPA-Human LF Sugar minus	napktfqfindqikfiins

1 qatateqeea ecatqqatqe aatqaaqaqa qqqetetqet qtqtqctqet qctqtqtqqa md amkr qlc cvl llcq 61 gcagtetteg tttegeeeag egeeggeggg catggggaeg ttggeatgea tgtgaaagaa avf vsp sagg hgd vg m hvke 121 aaggagaaaa acaaggacga aaacaagcgt aaagacgaag aacgtaataa aacacaggag keknkdenkr kde ern kt qe 181 gaacacttaa aggagatcat gaagcacata gtaaagattg aggtaaaagg cgaagaggct ehl keim khi v ki ev k qee a 241 qtaaaqaaqq aqqcaqcaqa aaaactgttq qaqaaqqtqc cttctqacqt cttaqaqatq vkkeaaekllekvpsdvlem 301 tataaggeea teggeggtaa gatetatate gtggacggag acateactaa acacatatet y kaigg kiyi v dg dit khis 361 ctcgaagctc tctccgagga caagaaaaag attaaagaca tctacgggaa ggatgcctta lealse dkkk ikd i yg kdal 421 ttgcacgage actacgttta cgcaaaggag ggctatgage ccgtqctcqt tattcagagt lhehyvyakegyepvl viqs 481 agtgaggact acgtcgagaa taccgagaaa gctctgaatg tgtattacga gatcggaaag sedy ventek aln vyyei gk 541 attotatoco aggacatoet atccaaaato aaccaaccat accaaaatt cottaatatt ils rdilskingpyqk fldv 601 cttaacacaa tcaaaaacgc qtcaqatagc qacgqqcagq atcttctqtt tacaaatcaa lntikn asds dgq dll ftnq 661 ctcaaqqaac accccactga tttcagcgtq qaqttcctcg aqcaqaattc taacqaaqtc lke hpt dfsveflegn snev 721 caggaggtgt tegecaagge atttgegtae tatategaae eecageateg egatgtgete

q'e'v fa k'a fa y y i'e p q h'r'd'v 1 781 cagotgtacg ccccggaggc atttaactac atggacaaat tcaatgaaca ggagattaat

1

q l y a p e a f n y m d k f n e q e i n 841 ctgtctctgg aggaactgaa agaccagtga ggatcc l s l e e l k d q -

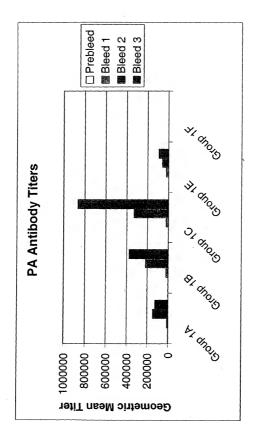


FIG. 15A

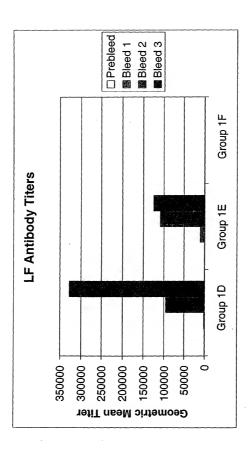


FIG. 15B

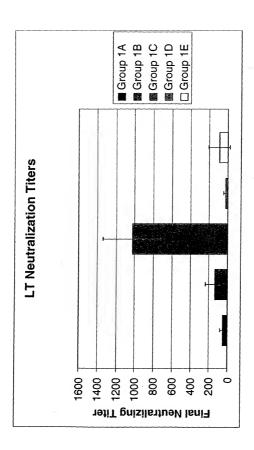


FIG. 15C

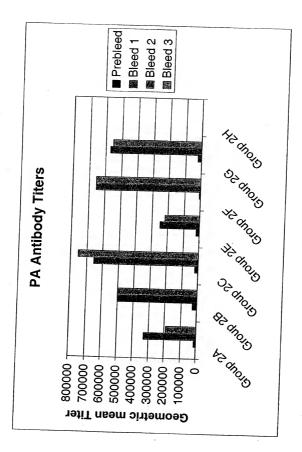


FIG. 16A

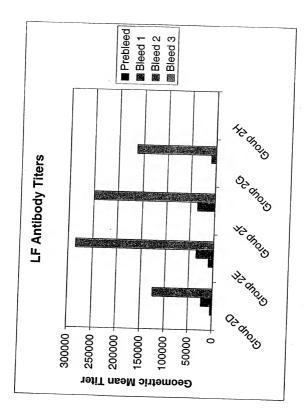


FIG. 16B

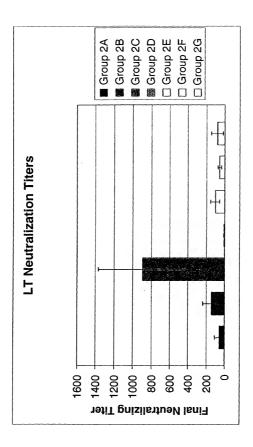


FIG. 16C

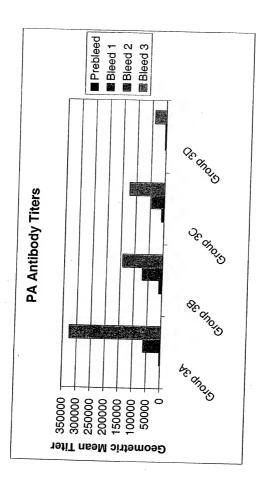


FIG. 17A

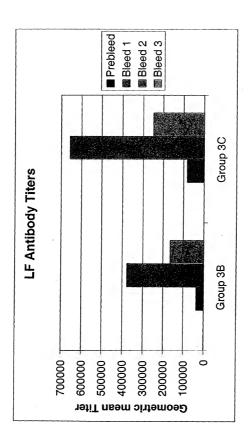


FIG. 17B

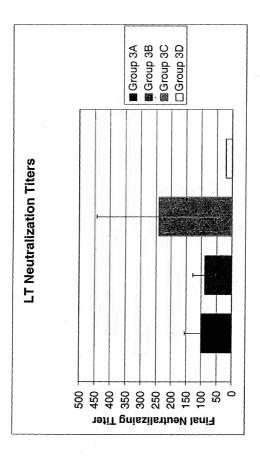


FIG. 17C

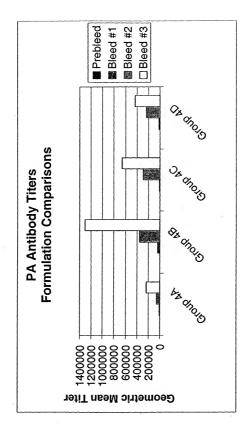


FIG. 18A

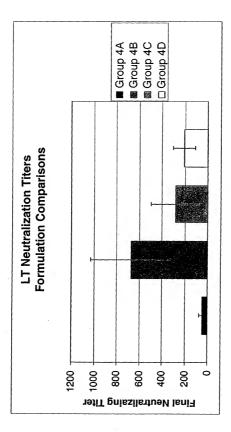


FIG. 18B

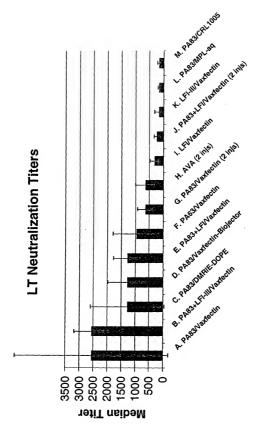


FIG. 1

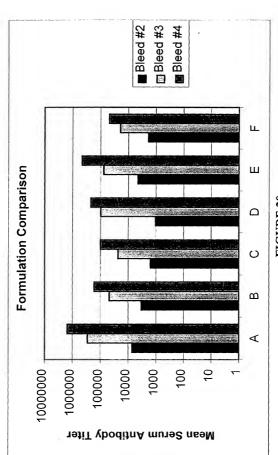


FIGURE 20

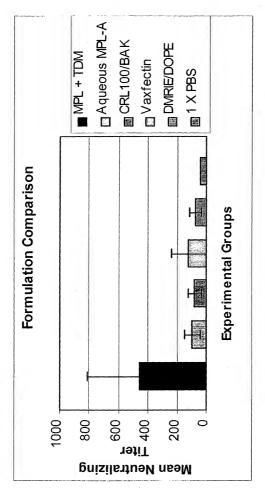


FIGURE 21